

うつキーワードと感情変動に基づくブログからのうつ検出手法

松本 和幸* 吉岡 伸浩† 北 研二* 任 福継*

*徳島大学大学院ソシオテクノサイエンス研究部 †三協商事株式会社

{matumoto, kita, ren}@is.tokushima-u.ac.jp

1 はじめに

日本におけるうつ病や躁鬱病の患者数は、年々増加傾向にある。日常のストレス等からうつ病にかかってしまい、物質的・金銭的に余裕のある人が自殺してしまうケースも増えている。その大きな原因は、我々の心の豊かさが低下してきていることであると思われる[1]。ここでの心の豊かな状態とは、喜びや悲しみ、怒り、などといった感情を適切に理解し、表現し、自ら実感できるような状態を表す。うつ病になってしまうと、通常なら快感情が生起するような場合でも、不快感情が生起するなど、悲観的思考に陥りやすい。また、情緒が不安定になることも、うつ病の症状のひとつとされる。こうした症状は、他の病気に比べ、変化を見落としやすいため、気付いたときには進行していて、手遅れになっていることもよくある。従来、うつ病などの精神疾患には、専門の精神科医による問診によって判断する方法や、DSM-IV-TR[2]という、質問項目に回答することで診断できる手引書の利用などがあった。しかし、こうしたうつ病診断は本人が自覚していない限り、受診することがないことから、うつ病の発見が遅れてしまう。

我々は、ブログの書き手に、うつ傾向があることを自覚させることにより、うつ病の早期発見・早期治療に役立てることができると考えた。まず、うつ傾向の有無を、ブログ内容から推定することができるかをテキスト分類手法を用いた実験により明らかにし、どのような特徴がうつ判定に有効であるかを調査する。また、人間の心の変化をとらえるための特徴として、時間経過に伴う記述内容の変化から読み取れる感情状態の変化がある。多くのブログが、毎日の出来事や心情を時系列順に書き記したものであるため、そこから時間的変化を得ることが出来る。本稿では、うつ特有の言葉遣いを収集したうつキーワード辞書に基づくうつ検出手法と、感情状態の変化に基づいたうつ検出手法を提案する。

2 うつ傾向のブログの収集と解析

2.1 うつ・非うつ傾向の比較

うつ傾向のある著者のブログを大量に収集し、それをうつ傾向検出のための基本となるデータとして用いることを考える。うつ傾向の有無は、人手によりブログを見て判断を行う。うつ傾向の判定指標として、まず、著者自身がうつ病であることをブログ上で公言していることがある。このような著者は、うつ傾向があることを自覚していて、通院経験があったりもする。一方で、うつ傾

表 1: 出現頻度上位 10 件の共通素性: 双方における出現頻度 (うつ:非うつ)

e	思う (507:286), 涙 (108:51), 楽 (94:30), 泣く (84:75), 好き (79:90), 辛い (73:12), 不安 (70:9), 心配 (68:13), 痛み (62:20), 怒る (52:16)
d	重要・大切・根底 (864:429), 正解・適当・確実 (639:235), 適当・適合・適期 (575:281), 能才 (479:186), 粗悪・不備・素人 (459:287)

向であることを自覚していない著者がいたりもする。一般に、こうした著者のブログを見つけ出すことは難しい。そのため、うつ傾向があると明確に公言している著者 30 名のブログ (うつ傾向ブログ) と、そうでない著者 30 名のブログ (非うつ傾向ブログ) に分類する。60 名の各著者ごとに、約 30 日分、述べ 1800 記事を収集した。そして、収集したうつ傾向ブログと非うつ傾向ブログとの間に何らかの差があるかどうかを調べることにした。

それぞれに使用されている単語の統計を行なったところ、幾つかの点で共通点が見られた。ここでは、感情表現辞書と日本語アブレイザル評価表現辞書 (JAppraisal 辞書) [3] に基づき抽出した素性を分析対象とした。感情表現辞書は、中村の「感情表現辞典」[4] において 10 種類の感情に分類された表現のうち、単語のみを登録した辞書である。「日本語アブレイザル評価表現辞書-態度評価編」は、国立国語研究所が開発した「愛する」「非道」などの評価表現を、肯定的か否定的か (評価極性) だけでなく、評価基準の種類によって詳細に分類し、体系化した辞書である。登録語彙数は、約 8000 語である。

表 1 に、うつ傾向ブログと非うつブログの両方に共通して出現した素性の一部を示す。また、表 2 に各分類一方のみに出現した素性の一部を示す。表中の e は、感情表現, d は、JAppraisal 辞書の “delicatefeature” を示す。

この分析結果からみて、単純な感情表現については、うつ傾向の有無に関わらず共通して多く用いられる。一方で、うつ傾向の有無を特徴付けている、うつ/非うつそれぞれに偏って出現する語句に関しては、著者の用いる表現の偏りなども考慮しないといけないため、この結果のみからは単純に判断できない。しかし、うつ傾向にのみ出現する感情表現については、「苦しむ、倦怠、不快」のように、明らかにネガティブな語が多く含まれている。一方で、非うつブログのみに出現する感情表現にもネガティブな語があるが、「恋、愉快」などといったポジティブな語も含まれている。このことから、うつ傾向の判定を単語に基づいて行えらと考えた。

表 2: 一方にだけ出現した素性の比較

	うつ	非うつ
e	苦しむ (24), 倦怠 (22), 不快 (12), 無気力 (11), 嫌気 (10), 悲観 (10), 鬱ぐ (9), 苦しみ (9), 動悸 (9), ほっと (8)	鳥肌 (8), やれやれ (6), 寒気 (6), 案外 (5), 恋 (4), 心外 (3), 煩い (2), 緊迫 (2), 自棄 (2), 心地好い (2), 愉快 (2)
d	憐れむ (12), 安心・安堵 (9), 気楽さ・気軽さ (7), 悶着・仇・疎外 (6), 億劫・躊躇い (4)	恩賞・報恩・思いやり (10), 邪魔・束縛・責任転嫁 (4), 陰気・汚らしさ (3), 期待・信頼・信用 (3), 嫌う・疎む/恨み・憎しみ (3)

表 3: Bag-of-Words による実験結果

	NB		SVM		MEM	
	(10)	(60)	(10)	(60)	(10)	(60)
うつ	92.4	88.9	84.7	13.0	91.9	66.6
非うつ	83.8	76.0	72.7	29.9	59.3	48.4
平均	88.1	82.4	80.8	21.2	75.6	66.6

2.2 Bag-of-Words によるうつ傾向判別実験

テキスト分類によく用いられる手法として、文書に含まれる単語を特徴とするものがあり、Bag-of-Words(BOW)と呼ばれる。この方法は、単純ではあるが、強力な分類手法として、ベースライン手法としてよく用いられる。本研究では、うつ傾向があるか否かを2値分類するため、分類器として、スパムメールの分類に有効な分類器として知られるナイーブベイズ分類器(NB)とパターン認識でよく用いられるサポートベクトルマシン¹(SVM)、また、テキスト分類タスクで使われることが多い最大エントロピーモデル [5](MEM)を用いる。これらの手法をベースライン手法として、うつ傾向のある著者のブログとそうでない著者のブログを分類するモデルを各手法において構築し、うつ傾向判定の予備実験を行なった。形態素解析には、日本語形態素解析器 MeCab²をデフォルト設定で、辞書には IPA-dic を用いた。SVM はデフォルトの線形カーネルを用い、ハイパーパラメータもデフォルトの値を用いた。まず、10 分割交差検定を行なった結果と、ある著者とそれ以外の著者のブログとでテストデータと訓練データに分割し、それを繰り返すことで 60 分割交差検定を行なった結果を、表 3 に示す。表中の (10), (60) は、それぞれ 10 分割、60 分割交差検定の結果を示す。この結果は、正解した記事の数を実験に用いたすべての文の数で割り、正解率を計算したものである。

記事単位での実験と、著者単位での実験のどちらにおいても、NB と比較して SVM はあまり良い結果が出ていない。これは、過学習が1つの原因と考えられるため、過学習を避けるためにハイパーパラメータを調整すれば多少正解率が高くなると思われる。

¹<http://chasen.org/~taku/software/TinySVM/>

²<http://mecab.sourceforge.net/>

2.3 感情素性を用いたうつ傾向判別実験

前節での予備実験の結果、BOW により、ある程度高精度のうつ判定が行えた。しかし、訓練データに含まれない単語がテストデータに含まれることが、判定失敗の原因になると考えられるため、単語そのものではなく、単語の属性などを素性として用いることを考える。単語の素性として、感情極性対応表 [6] から得た感情極性値、また、感情表現辞書における単語の感情、JAppraisal 辞書における属性を用いる。JAppraisal 辞書は、感情や態度を表した表現を登録したものであるため、辞書中のすべての語がポジティブかネガティブのどちらかに分類されている。この辞書における、“delicatefeature” という属性は、分類体系の最下位カテゴリである。この属性を用いることで、より詳細な違いを見分けることができると考えられるため、素性に含める。これらの素性の有効性確認のため、NB, SVM, MEM を用いて、各素性を使ったうつ判定実験を行なった。実験は、前節で用いたブログ記事を対象に著者単位での交差検定 (60-fold) で行なった。その結果を表 4 に示す。

表 4: 各素性を用いた実験結果 (%)

素性	正解率 (%)		
	NB	SVM	MEM
感情極性値	69.0	25.5	52.5
感情表現	63.9	21.9	56.2
単語の感情	57.5	51.0	55.5
delicatefeature	60.9	29.2	51.1

単純な BOW の方が比較的分類精度が高い。これは、感情表現や評価表現以外の表現も、うつ傾向を検出する手がかりとして重要であることを示唆するものである。また、NB を用い、感情極性値を素性として用いた場合に、最も正解率が高くなった理由として、感情極性対応表のエントリ語数が3つの辞書の中で最大であることも関連していると考えられる。

3 うつキーワード辞書の構築

前節では、機械学習に用いる素性を複数種類試すことにより、うつ傾向の判別を試みた。しかし、結果から言うところ、単純に形態素の出現頻度を学習させたほうが高精度でうつの判別が可能である。原因として、感情表現等の特徴のみではうつの特徴としては不十分であることが考えられる。本研究では、うつ特有の特徴を考慮するため、精神医学の分野でうつ病判定基準として用いられる DSM-IV-TR をもとに、うつ病の特徴を表す語句 (うつキーワード) を手作業で収集し、辞書を作成した。

作成したうつキーワード辞書に含まれるキーワードの分類とその数について、表 5 に示す。

ここで、うつキーワードを素性とすることで、他の素性と比較してどの程度有効であるか検証するため、単語の感情とうつキーワード (完全一致で抽出) を素性とした場合の、NB によるうつ傾向判定実験を行った結果、表 6 のような結果が得られた。うつキーワード

表 5: うつキーワード辞書

身体症状	語句数
不眠または睡眠, 頭重感, 頭痛, めまい	179
食欲不振または食欲旺盛, 胃部不快感, etc.	218
肩こり, 背中や腰など体の痛み	124
息苦しい, 動機がする	150
手足の痺れ感, 嫌な汗や寝汗が出る	150
排尿困難, 性欲低下, 異性への関心	96
精神症状	語句数
物事をやるのがおっくうで早くできない	72
集中力が落ち, 仕事を能率よくできない	116
人に会いたくない, 人と一緒にいたくない	86
興味や喜びの喪失	105
自殺への思い	77
憂鬱な気分 (不安や焦燥感からのイライラ)	173

を含めた場合のほうが高い正解率を得られたことから, うつキーワードがある程度有効であることがわかる。ただ, 感情極性値を用いた場合より正解率が低いため, キーフレーズの照合方法の検討などが必要と考えられる。

表 6: うつキーワードを追加した場合の結果 (60-fold)

うつ	非うつ	平均
69.7	61.1	66.0

4 感情変動に基づくうつ傾向検出

一般に, 時々刻々と変化する感情をリアルタイムで検出することは難しい。特に, 一般的なブログは, 感情の生じた時間とそれについて書かれた記事を投稿する時間に時差が生じるため, 即時性という面で, チャットやマイクロブログ等よりも幾分か劣る。しかし, Twitter に代表されるような一言メッセージによるマイクロブログと比較して, 一つの記事から得られる情報量が多いため, より確実性の高い解析が行えることもまた事実である。

Scott[7] は, Twitter での発言内に含まれるポジティブ/ネガティブを表す単語の出現傾向を分析し, 感情の変化と時間/季節との関連性を, 世界各国の Twitter から取得した 5 億以上のメッセージをもとに導き出し, 文化差を超えて感情変動に共通性があることを示した。ブログからうつ傾向を検出するためには, 一般的な感情の変動と逸脱するような感情変動に着目する必要があると考えられる。うつ傾向があるとされる人の特徴として, 感情の変化に乏しかったり, 感情の変化が著しかったりと, 情緒不安定な傾向がある。うつ傾向にある人の書いた文章と, そうでない人の文章とで, 感情変化に違いがあると考えられる。

2 節で説明したブログ記事データセットとは別に, うつ傾向のあるブログ著者 10 名, うつ傾向の無いブログ著者 10 名の計 20 名分のブログを, 約 1ヶ月分 (30 日分) ずつ収集した。また, それに対して人手により, 各ブログ記事が表している感情とその程度を 0 から 1 までの実数値で付与した。このデータセットを用い, 感情の変動を分析した。ここで付与した感情の種類は (喜び,

怒り, 悲しみ, 期待, 焦り, 憎しみ, 驚き, 好き) の 8 感情である。

まず, うつと非うつとで, 各感情の値がどのように異なるかを平均値を計算し比較してみたところ, 図 1 のようになった。「喜び」と「悲しみ」は, うつと非うつとで, 平均値に明確な差が出ている。うつの場合「悲しみ」が強く出る傾向があり, 非うつの方では「喜び」が強く出る傾向があることが分かる。

この特徴を生かし, 1ヶ月間のブログにおいて「悲しみ」が「喜び」の 1.5 倍以上出現する場合にうつ傾向があると判定する。それ以外は, すべて非うつと判定する。実験の結果, うつ判定正解率 85% が得られた。

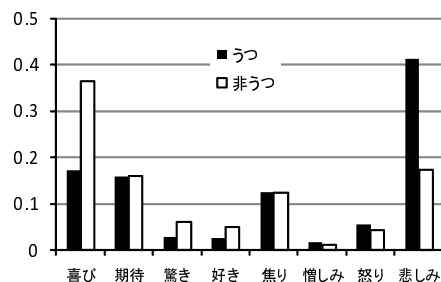


図 1: うつ/非うつの感情の値の平均値

次に, これらの感情の, 時系列的な変化について調べる。ポジティブ, ネガティブの感情の前日からの変化を, 以下に示す 6 種類のパターン (感情変動パターン) に分類した。

- F_1 ネガティブ感情が変化せず持続
- F_2 ネガティブ感情が増加
- F_3 ポジティブ感情が減少
- F_4 ポジティブ感情が変化せず持続
- F_5 ポジティブ感情が増加
- F_6 ネガティブ感情が減少

1 つのブログ記事に複数の感情の付与を許可しているため, 1 日分の変化は, 複数のパターンに当てはまる場合もある。このパターン分析を, 約 1ヶ月間分, 連続する各ブログ間において行い, 各著者単位で, F_1 から F_6 までのパターンの出現頻度を正規化して感情変動ベクトルを作成した。このベクトルに対し, 著者 i 以外の著者 j のブログから作成した感情変動ベクトルとの間の余弦類似度を基に, 類似度の高い順に k_{max} 位まで出力した際の, うつ/非うつのラベルの一致に関して平均適合率を求めた。まず, F_1 から F_6 まですべてを用いた場合の結果を, 図 2 に示す。横軸が k_{max} の値, 縦軸が平均適合率 (%) を表す。また, うつ傾向があると, ネガティブ感情が増加しやすいという点を考慮して, F_1, F_2, F_3 の 3 つのパターンのみを用いた場合の実験も行った。図 3 にその結果を示す。

次に, ポジティブ感情の考慮が, 誤って非うつと分類されてしまう原因とも考えられるため, F_1, F_2 のみを用いて実験を行った。その結果を, 図 4 に示す。

次に, 感情変動のパターン分けを, その変化の程度を条件として行うことを考える。同じ感情変動パターンに

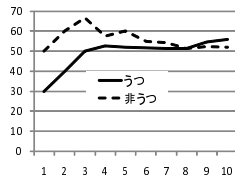


図 2: 実験結果 $F_1 \sim F_6$

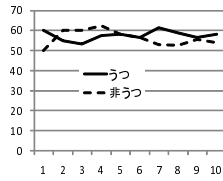


図 3: 実験結果 F_1, F_2, F_3

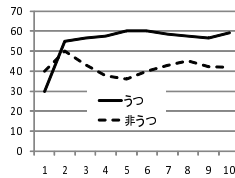


図 4: 実験結果 F_1, F_2

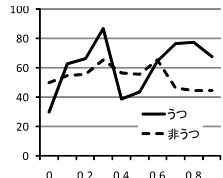


図 5: 実験結果 (閾値有)

分類されていても、変化が少ないものと多いものとは、異なる特徴として扱ったほうが良いと考えられるからである。以下に、パターン分けの規則の一部を示す。 NA_t は、感情変動計算対象の時間 t のブログにおける、ネガティブ感情の平均値を示す。また、 NA_{t-1} は、 NA_t より 1 つ前のブログにおけるネガティブ感情の平均値を示す。もし、これらの差の絶対値 $|NA_t - NA_{t-1}|$ が閾値 $threshold$ より大きければ、ネガティブ感情が増加したと見なし、 t におけるパターン ptn_t は、 F_2 、または F_6 となる。閾値 $threshold$ 以下であれば、変化しなかったものと見なし、 F_1 となる。ポジティブ感情とネガティブ感情に対し、このようなパターン分け規則をそれぞれ作成し、実験した結果を、図 5 に示す。横軸が、 $threshold$ の値を示す。 $F_1 \sim F_6$ すべてを用い、 k_{max} の値を 1 とした。また、 $0 \leq threshold \leq 1.0$ とする。

```

if  $|NA_t - NA_{t-1}| > threshold$  then
  if  $NA_t - NA_{t-1} > 0$  then
     $ptn_t \leftarrow F_2$ 
  else if  $NA_t - NA_{t-1} < 0$  then
     $ptn_t \leftarrow F_6$ 
  end if
else
   $ptn_t \leftarrow F_1$ 
end if

```

これらの結果から、感情変動ベクトルの類似度によるうつ傾向の検出には、 F_1, F_2, F_3 のパターンが他よりも有効であることが分かる。一方、 F_1, F_2 のみを用いた場合には、非うつ傾向の判別に失敗しやすくなる。また、パターン分け規則を用いた場合には、最大で 86.6% のうつ判定正解率が得られた。このとき、非うつ判定率は、65.6% であった。すべてのパターンを用い、差の閾値を導入したことにより、感情変動の細かな違いを識別できるようになり、正解率が向上したと考えられる。しかし、学習データとしてうつ/非うつ各 10 名分のみを用いただけであり、この検出手法では、各ブログ記事への感情タグの付与が必要であることから、これらの特徴を用いるよりも、ポジティブ/ネガティブ素性を利用した方が効

率が良いと考えられる。

5 おわりに

本稿では、うつブログの収集と解析を行なった結果をもとに、ブログに書かれた内容およびブログ記事の感情の変化からうつ傾向の検出を行う手法を提案した。評価実験の結果、感情変動ベクトルを用いた手法は、単にポジティブ/ネガティブ感情の増加/減少をパターン化するだけでなく、その変化の度合いを考慮して一定の条件を加えることで、うつ検出精度が向上するということが分かった。

今後は、うつ傾向の検出精度を高める手法を検討すると共に、吉岡ら [8] の提案した理論に基づいて、うつ傾向が検出された著者の心の状態を良い方向へ向かわせるための癒しの方法として「喜び」の原因となる物事を自動的に抽出する手法を提案したい。

謝辞

本研究の一部は、科学研究費補助金（挑戦的萌芽研究：21650030）により実施した。

参考文献

- [1] 任福継. 豊心工学. 電子情報通信学会技術研究報告 TL, 思考と言語, Vol. 110, No. 244, pp. 19–24, 2010.
- [2] American Psychiatric Association. DSM-IV-TR 精神疾患の分類と診断の手引き. 医学書院, 2006.
- [3] 国立国語研究所コーパス開発センター. 日本語アブレザル評価表現辞書 (JAppraisal 辞書) ~ 態度評価編 ~ Ver.1.0 仕様説明書. 2011.
- [4] 中村明. 感情表現辞典. 東京堂出版, 1993.
- [5] Adwait R. A simple introduction to maximum entropy models for natural language processing. *Technical Report*, Vol. 97, No. 08, 1997.
- [6] 高村大也, 乾孝司, 奥村学. スピンモデルによる単語の感情極性抽出 (自然言語). 情報処理学会論文誌, Vol. 47, No. 2, pp. 627–637, 2006.
- [7] Scott A. G. and Michael W. M. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, Vol. 333, No. 6051, pp. 1878–1881, 2011.
- [8] 吉岡伸浩, 松本和幸, 任福継. Blog 情報分析による作者の喜びの原因検出について. 電気学会電子・情報・システム部門大会論文集, pp. 1446–1448, 2009.